

University of Groningen

Analysis and Control of Strategic Interactions in Finite Heterogeneous Populations under Best-Response Update Rule

Ramazi, Pouria; Cao, Ming

Published in:

Proc. of 2015 IEEE 54th Annual Conference on Decision and Control (CDC)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2015

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Ramazi, P., & Cao, M. (2015). Analysis and Control of Strategic Interactions in Finite Heterogeneous Populations under Best-Response Update Rule. In *Proc. of 2015 IEEE 54th Annual Conference on Decision and Control (CDC)* (pp. 4537). IEEE (The Institute of Electrical and Electronics Engineers).

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Analysis and Control of Strategic Interactions in Finite Heterogeneous Populations under Best-Response Update Rule

Pouria Ramazi and Ming Cao

Abstract—For a finite, well-mixed population of heterogeneous agents playing evolutionary games choosing to cooperate or defect in each round of the game, we investigate, when agents update their strategies in each round using the myopic best-response rule, how the number of cooperating agents changes over time and demonstrate how to control that number by changing the agents' payoff matrices. The agents are heterogeneous in that their payoff matrices may differ from one another; we focus on the specific case when the payoff matrices, fixed throughout the evolution, correspond to prisoner's dilemma or snowdrift games. To carry out stability analysis, we identify the system's absorbing states when taking the number of cooperating agents as a random variable of interest. It is proven that when all the agents update frequently enough, the reachable final states are completely determined by the available types of payoff matrices. As a further step, we show how to control the final state by changing at the beginning of the evolution, the types of the payoff matrices of a group of agents.

I. INTRODUCTION

For over three decades, biologists, economists and sociologists have utilized evolutionary game theory to understand how a certain level of cooperation can be reached among selfish interacting individuals [1][2]. Such evolutionary game theoretic models have proven to be promising in engineering network systems as well especially in the context of growing applications of networks of autonomous agents performing group tasks [3][4][5]. While simulation results have suggested there are effective mechanisms to explain the emergence and promotion of cooperation among self-interested agents [6][7], rigorous mathematical statements with proofs are still in great demand. Moreover, the existing proven mathematical results, e.g., [8][9][10], usually assume at least one of the followings: (i) an infinite population under continuous-time dynamics, (ii) deterministic dynamics, and (3) homogeneous agents. While these assumptions help to simplify the mathematical setup and gain insight into the theoretical aspects of cooperation mechanisms, an interesting research line that has not been explored to its full potential is to study a finite heterogeneous population of game-playing agents under stochastic discrete-time updating dynamics.[7]

Interesting simulation results have been conducted for heterogeneous populations in [11], [12] and [13] where the agents are associated with different payoff matrices. Some mathematical statements have been provided in [14] and [15]

for the myopic best response update rule [15] but only for the case when the population is homogeneous and the game is symmetric. Some researchers have studied the stochastic stability of different strategies in the population when the update rule is noisy [16] [14]. Besides the results addressing the stability issue of strategically interacting populations, there is an emerging trend to investigate how to control such populations. Different methods have been suggested to control the number of agents with a specific strategy in a decision-making population. For example, in [17][18], the strategies of a group of agents are fixed to a desired strategy for the duration of the game, in [19], the payoffs of a stochastic snowdrift game are changed and in [20], some changes in terms of an emission tax rate and the price of an emission permit are applied to the payoffs of the agents.

In this paper, we continue our recent work [21] by considering a finite number of agents, each belonging to a type to which a (possibly unique) payoff matrix is associated. At each time an agent is randomly chosen from the population to update her strategy that is either to cooperate or defect, according to the myopic best response update rule. In such a population, we explicitly determine the number of cooperators of each type after a sufficiently long time, find the absorbing states, and clarify their stability properties. Consequently, we show that based on the distribution of the types in the population, the total number of cooperators in the long run either reaches a certain steady state or fluctuates between two constants. After presenting the stability analysis, we investigate whether it is possible to control the number of cooperators in the long run by changing the types of the agents. Specifically, for a given population and types, we find how to change the types of the agents to have a desired number of cooperators in the long run.

The rest of the paper is organized as follows. In Section II, we construct the framework of our work and introduce the population game that the agents are involved in. In Section III, we provide the main stability result of this paper, and explain the evolution of the number of cooperators in each type of agent. In Section IV, we focus on controlling the total number of cooperators in the population after a sufficiently long time. It is shown how to achieve a desired reachable number of cooperators in the long run by changing the types of a group of agents.

II. POPULATION STRATEGIC DYNAMICS

We consider a finite, well-mixed population of n agents that are participating in a population game evolving over time $t = 0, 1, \dots$. Each agent can choose either to cooperate

The work was supported in part by the European Research Council (ERC-StG-307207).

P. Ramazi and M. Cao are with ENTEG, Faculty of Mathematics and Natural Sciences, University of Groningen, The Netherlands, {p.ramazi, m.cao}@rug.nl

(C) or defect (D). At each time t , an agent is randomly activated to update her strategy according to how well she is doing when she plays her current strategy against the average population. More specifically, the four possible payoffs of an agent i , $i = 1, \dots, n$, are summarized in the 2×2 payoff matrix A_i

$$A_i = \begin{matrix} & \begin{matrix} C & D \end{matrix} \\ \begin{matrix} C \\ D \end{matrix} & \begin{pmatrix} R_i & S_i \\ T_i & P_i \end{pmatrix} \end{matrix}, \quad (1)$$

where the payoffs R_i, T_i, S_i and P_i are real numbers corresponding to strategy pairs C -against- C , D -against- C , C -against- D and D -against- D respectively. Let $s_i(t)$ denote the agent i 's strategy at time t , which is defined to be $[1 \ 0]^T$ when agent i chooses to cooperate at t and defined to be $1 - [1 \ 0]^T = [0 \ 1]^T$ otherwise, with $\mathbf{1} = [1 \ 1]^T$. Let $x_C(t)$ denote the proportion of cooperators in the whole population of n agents at time t and define the *average population vector* $s_C(t) = [x_C(t) \ 1 - x_C(t)]^T$. Then agent i 's payoff at time t against the average population is calculated by

$$u_i(s_i(t), s_C(t)) = s_i(t)^T A_i s_C(t).$$

The myopic best-response strategy update rule for agent i dictates that agent i sticks to her current strategy if her alternative strategy does not give her higher payoff, and otherwise she switches her strategy, namely

$$s_i(t+1) = \begin{cases} s_i(t) & \text{if } u_i(s_i(t), s_C(t)) \geq u_i(\mathbf{1} - s_i(t), s_C(t)) \\ \mathbf{1} - s_i(t) & \text{otherwise} \end{cases}. \quad (2)$$

Obviously, how agent i calculates her payoff, or more precisely, the exact form of her payoff matrix A_i in (1), is critical for the strategy update dynamics. In this paper, we focus on those A_i with special structures. We require that the entries of A_i satisfy that

$$T_i > R_i > \max\{S_i, P_i\}. \quad (3)$$

Payoff matrices with this property correspond to either a prisoner's dilemma (PD) or snowdrift (SD) game. Following the setup in [21], we assume that all the A_i correspond to either a specific PD game satisfying $P_i > S_i$ or correspond to possible variations of the SD game satisfying $S_i > P_i$. As it is later shown in Lemma 1, the comparison of the coefficient $\frac{S_i - P_i}{T_i - R_i + S_i - P_i}$ with the ratio of cooperators determines whether agent i changes her strategy when she is chosen to update. Correspondingly, all SD payoff matrices can be classified into $l > 0$ types according to their different values of the quotients $\frac{S_i - P_i}{T_i - R_i + S_i - P_i}$ which have to lie strictly in $(0, 1)$. We call all those agents with the PD payoff matrix the *PD agents*. We label all those l types of SD payoff matrices by $1, \dots, l$ according to the *descending* order of magnitude of the quotients $\frac{S_i - P_i}{T_i - R_i + S_i - P_i}$, and call all those agents with the j th, $j \in \{1, \dots, l\}$, type SD payoff matrix the *SD_j agents*.

Then there are altogether $l + 1$ types among the n agents. Let n_{PD} denote the number of PD agents and n_{SD_j} , $j = 1, \dots, l$, the number of SD_j agents. Then given the initial conditions of all the agents with their fixed types and initial

strategies, under the best-response strategy update rule (2), the *type population* defined by

$$p = (n_{SD_1}, n_{SD_2}, \dots, n_{SD_l}, n_{PD})$$

evolves in the set

$$\mathcal{P}_n := \left\{ p \in \mathbb{Z}_+^{l+1} \mid 0 \leq p_i \leq n, \sum_{i=1}^{l+1} p_i = n \right\}.$$

Given the types SD_1, \dots, SD_l and PD , the type population $p \in \mathcal{P}_n$ and under the update rule (2), denote the number of cooperators in the population at time t by $n_C(p, t)$. We simplify this notation to $n_C(t)$ in Sec. III, when the type population p is clear. Since we want to know the trends of the changes of the number of cooperators in the population, we denote the distributions of the cooperators in different types by the *population state*

$$a(t) := (n_{SD_1}^C(t), n_{SD_2}^C(t), \dots, n_{SD_l}^C(t), n_{PD}^C(t)),$$

where each component of the vector is the number of cooperators in the corresponding type.

In the next section, we first study how the population state evolves over time and in particular, what its absorbing states are if there is any.

III. STABILITY ANALYSIS OF THE POPULATION GAME

We first describe what the absorbing states of the population state $a(t)$ look like and then prove that indeed they are the absorbing states. For the types SD_1, \dots, SD_l and PD and a type population $p \in \mathcal{P}_n$, define

$$k_p := \begin{cases} \arg \min_{k \in \mathcal{L}} \sum_{i=1}^k p_i \geq n_{SD_{k_p}}^* & \sum_{i=1}^l p_i \geq n_{SD_l}^* \\ l + 1 & \sum_{i=1}^l p_i < n_{SD_l}^* \end{cases}$$

where $\mathcal{L} = \{1, \dots, l\}$, and for $j \in \mathcal{L}$, $n_{SD_j}^*$ is defined to be the value of $n \frac{S_j - P_j}{T_j - R_j + S_j - P_j}$ for the payoff matrix of the SD_j agents. We assume that the following holds

$$\lfloor n_{SD_i}^* \rfloor > \lceil n_{SD_{i+1}}^* \rceil \quad \forall i \in \{1, \dots, l-1\}. \quad (4)$$

Define the following $(l+1)$ -dimensional row vectors

$$\begin{aligned} a^* &:= (n_{SD_1}, \dots, n_{SD_{k_p-1}}, 0, \dots, 0), \\ a_-^* &:= (n_{SD_1}, \dots, n_{SD_{k_p-1}}, \lfloor n_{SD_{k_p}}^* \rfloor - \sum_{j=1}^{k_p-1} n_{SD_j}, 0, \dots, 0), \\ a_+^* &:= (n_{SD_1}, \dots, n_{SD_{k_p-1}}, \lceil n_{SD_{k_p}}^* \rceil - \sum_{j=1}^{k_p-1} n_{SD_j}, 0, \dots, 0). \end{aligned}$$

Then define the set \mathcal{A} as follows. If $k_p \neq l + 1$,

$$\mathcal{A} := \begin{cases} \{a^*\} & \sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SD_{k_p}}^* \\ \{a_-^*, a_+^*\} & \sum_{j=1}^{k_p-1} n_{SD_j} < n_{SD_{k_p}}^* \end{cases}$$

and if $k_p = l+1$, then $\mathcal{A} := \{a^*\}$. With the above definitions in hand, we are ready to present the main stability result of this section.

Theorem 1: Consider the types SD_1, \dots, SD_l and PD and the type population $p \in \mathcal{P}_n$. When every agent is activated independently with a nontrivial probability, under the update rule (2) and for any initial population state $a(0)$, there exists some time τ such that almost surely

$$a(t) \in \mathcal{A} \quad \forall t \geq \tau. \quad (5)$$

Note that the theorem implies that the set \mathcal{A} is not only invariant but also globally asymptotically attractive. According to the theorem, there exists some time τ such that for all $t \geq \tau$, it holds that

$$\begin{cases} a(t) = a^* & k_p \neq l+1 \text{ and } \sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SD_{k_p}}^* \\ a(t) \in \{a_-^*, a_+^*\} & k_p \neq l+1 \text{ and } \sum_{j=1}^{k_p-1} n_{SD_j} < n_{SD_{k_p}}^* \\ a(t) = a^* & k_p = l+1 \end{cases}.$$

In other words, after a finite time, almost surely the population state reaches a situation where for $i \in \{1, \dots, k_p-1\}$ all of the SD_i agents cooperate, and for $i \in \{k_p+1, \dots, l\}$, all of the SD_i agents and also all of the PD agents defect. Moreover, for $k_p \neq l+1$, if $\sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SD_{k_p}}^*$, all of the SD_{k_p} agents defect. If $\sum_{j=1}^{k_p-1} n_{SD_j} < n_{SD_{k_p}}^*$, then in case $a_+^* = a_-^*$ or equivalently $\lfloor n_{SD_{k_p}}^* \rfloor = \lceil n_{SD_{k_p}}^* \rceil$, we have that $n_{SD_{k_p}}^* - \sum_{j=1}^{k_p-1} n_{SD_j}$ of the SD_{k_p} agents cooperate and the rest defect. However, in case $a_+^* \neq a_-^*$, then the number of SD_{k_p} agents who cooperate, belongs to the set $\{\lfloor n_{SD_{k_p}}^* \rfloor - \sum_{j=1}^{k_p-1} n_{SD_j}, \lceil n_{SD_{k_p}}^* \rceil - \sum_{j=1}^{k_p-1} n_{SD_j}\}$. It can be shown that the number actually fluctuates between the elements of this set, which is omitted here due to page limit.

The main idea of the proof of Theorem 1 is as follows according to a sequence of lemmas to be developed. Note that due to page limit, we omit the proofs of most of the lemmas. First the case $k_p \neq l+1$ is considered. In Lemma 2 it is shown that \mathcal{A} is an invariant set. Then through Lemmas 3 and 4 it is shown that if the number of cooperators in the population is greater or smaller than the number of cooperators indicated in \mathcal{A} , it will return to the number indicated in \mathcal{A} after a finite time. In addition, with the help of Lemmas 5 and 6 it is proven that all of the $SD_1, SD_2, \dots, SD_{k_p-1}$ agents will become cooperators after a finite number of time steps, and will not change their strategies afterwards. Similarly it can be shown that all of the $SD_{k_p+1}, SD_{k_p+2}, \dots, SD_l$ agents will become defectors after a finite time, and will not change their strategies afterwards. Finally in view of all the proven facts, one can prove that a reaches and remains in \mathcal{A} after a finite time. We skip the discussion for the case when $k_p = l+1$ due to page limit.

To carry out the proof, we need to determine the best response of each type of agents to the average population s_C . Define $\beta_i(s_C(t))$, the *best-reply(response)* for player i to the average population at time t , as the set of pure strategies

$k \in \{C, D\}$ such that no other pure strategy gives her a higher payoff against $s_C(t)$:

$$\begin{aligned} \beta_i(s_C) = \\ \{k \in \{C, D\} : u_i(k, s_C) \geq u_i(x, s_C) \quad \forall x \in \{C, D\}\} \end{aligned}$$

Clearly, the best reply of each agent is completely determined by her type. We use $\beta_X(s_C(t))$ to denote the best response of an X -type agent to the average population at time t .

Lemma 1: For any time $t \geq 0$, the best response of a PD agent is

$$\beta_{PD}(s_C(t)) = \{D\}$$

and the best response of an SD_j agent, $j \in \{1 \dots l\}$, is

$$\beta_{SD_j}(s_C(t)) = \begin{cases} \{C\} & n_C(t) < n_{SD_j}^* \\ \{C, D\} & n_C(t) = n_{SD_j}^* \\ \{D\} & n_C(t) > n_{SD_j}^* \end{cases}. \quad (6)$$

Proof: For the PD agents it is clear that the best response is always to defect according to their payoff matrix which is a PD payoff matrix. For an SD_j agent, in order to calculate the best reply, we have to find the pure strategy k such that $u_i(k, s_C)$ is maximized. From the definition of $u_i(\cdot, \cdot)$, this is equivalent to player i choosing the maximum row of the multiplication $A_i s_C$:

$$A_i s_C = A_i \begin{bmatrix} x_C \\ 1 - x_C \end{bmatrix} = \begin{bmatrix} (R_i - S_i)x_C + S_i \\ (T_i - P_i)x_C + P_i \end{bmatrix}. \quad (7)$$

Comparing the two entries of the final right-hand vector determines the best response

$$\beta_i(s_j) = \begin{cases} \{C\} & (R_i + P_i - T_i - S_i)x_C > P_i - S_i \\ \{C, D\} & (R_i + P_i - T_i - S_i)x_C = P_i - S_i \\ \{D\} & (R_i + P_i - T_i - S_i)x_C < P_i - S_i \end{cases}. \quad (8)$$

Now according to (3) and the fact that A_i is an SD payoff matrix, it holds that

$$\begin{aligned} \left. \begin{matrix} T_i > R_i \\ S_i > P_i \end{matrix} \right\} &\Rightarrow \begin{cases} R_i - T_i + P_i - S_i < 0 \\ P_i - S_i < 0 \end{cases}. \end{aligned}$$

Hence, in view of (8), we have that

$$\beta_{SD_j}(s_C) = \begin{cases} \{C\} & x_C < \frac{S_i - P_i}{T_i - R_i + S_i - P_i} \\ \{C, D\} & x_C = \frac{S_i - P_i}{T_i - R_i + S_i - P_i} \\ \{D\} & x_C > \frac{S_i - P_i}{T_i - R_i + S_i - P_i} \end{cases}.$$

Multiplying the conditions in n completes the proof. ■

As can be seen from the lemma, the best response of the PD agents is always to defect. So if any of these agents is chosen to update her strategy at some time, she updates to D and in case her strategy was already D in the previous time step, she sticks to it. Now since each agent is chosen with a nonzero probability at all time, after some finite time, almost surely all of the PD agents get the chance to update their strategies. Hence, after a finite time t_{PD} , almost surely (a.s.) all of the PD agents become defectors. Moreover, they do not change their strategies afterwards. Hence, a.s.

$$a(t) = (n_{SD_1}^C(t), n_{SD_2}^C(t), \dots, n_{SD_l}^C(t), 0) \quad \forall t \geq t_{PD}.$$

In the rest of this section, we consider the situation when all of the PD agents have already become defectors. That is, all time t are considered to be non-less than t_{PD} .

In the rest of the proofs we make use of the following fact. According to the definition of k_p , it holds that

$$\sum_{i=1}^j n_{SD_i} < n_{SD_j}^* \quad \forall j < k_p. \quad (9)$$

Now we claim that \mathcal{A} is an invariant set under the myopic best response update rule.

Lemma 2: Let $k_p \neq l+1$. If $\sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SD_{k_p}}^*$, then in case the population state vector $a(t)$ reaches a^* at some time t_r , it will remain there afterwards, i.e.,

$$(\exists t_r : a(t_r) = a^*) \Rightarrow (a(t) = a^* \quad \forall t \geq t_r). \quad (10)$$

On the other hand, if $\sum_{j=1}^{k_p-1} n_{SD_j} < n_{SD_{k_p}}^*$, then in case the population state $a(t)$ reaches the set $\{a_-^*, a_+^*\}$ at some time t_r , it will remain in the set afterwards, i.e.,

$$(\exists t_r : a(t_r) \in \{a_-^*, a_+^*\}) \Rightarrow (a(t) \in \{a_-^*, a_+^*\} \quad \forall t \geq t_r). \quad (11)$$

Proof. Let $\sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SD_{k_p}}^*$ be in force. According to the definition of a^* , it holds that

$$n_C(t_r) = \sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SD_{k_p}}^* \geq \lfloor n_{SD_{k_p}}^* \rfloor. \quad (12)$$

Now let $\sum_{j=1}^{k_p-1} n_{SD_j} < n_{SD_{k_p}}^*$ be in force. According to the definitions of a_-^* and a_+^* , it holds that

$$n_C(t_r) \geq \sum_{j=1}^{k_p-1} n_{SD_j} + (\lfloor n_{SD_{k_p}}^* \rfloor - \sum_{j=1}^{k_p-1} n_{SD_j}) \geq \lfloor n_{SD_{k_p}}^* \rfloor.$$

Hence, in any case $n_C(t_r) \geq \lfloor n_{SD_{k_p}}^* \rfloor$ holds. Hence, in view of (4),

$$n_C(t_r) > n_{SD_i}^* \quad \forall i > k_p \xrightarrow{(6)} \beta_{SD_i}(s_C(t_r)) = \{D\} \quad \forall i > k_p.$$

Hence, if any SD_i agent, $i > k_p$, is chosen at t_r while her strategy is D at t_r , she will not change her strategy at t_r+1 . On the other hand, all SD_i agents, $i > k_p$, were already defecting at t_r according to the definition of \mathcal{A} . Hence,

$$n_{SD_i}^C(t_r+1) = n_{SD_i}^C(t_r) = 0 \quad \forall i > k_p. \quad (13)$$

Now we consider the rest of the types. Let $\sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SD_{k_p}}^*$ be in force. According to the definition of a^* , it holds that

$$n_C(t_r) = \sum_{j=1}^{k_p-1} n_{SD_j} \stackrel{(9)}{<} n_{SD_{k_p-1}}.$$

Now let $\sum_{j=1}^{k_p-1} n_{SD_j} < n_{SD_{k_p}}^*$ be in force. According to the definitions of a_-^* and a_+^* , it holds that

$$\begin{aligned} n_C(t_r) &\leq \sum_{j=1}^{k_p-1} n_{SD_j} + (\lfloor n_{SD_{k_p}}^* \rfloor - \sum_{j=1}^{k_p-1} n_{SD_j}) \leq \lfloor n_{SD_{k_p}}^* \rfloor \\ &\stackrel{(4)}{\Rightarrow} n_C(t_r) < n_{SD_{k_p-1}}^*. \end{aligned}$$

Hence, in any case $n_C(t_r) < n_{SD_{k_p-1}}^*$ holds. Hence, according to (4),

$$n_C(t_r) < n_{SD_i}^* \quad \forall i < k_p \stackrel{(6)}{\Rightarrow} \beta_{SD_i}(s_C(t_r)) = \{C\} \quad \forall i < k_p.$$

Hence, if any SD_i agent, $i < k_p$, is chosen at t_r while her strategy is C at t_r , she will not change her strategy at t_r+1 . On the other hand, all SD_i agents, $i < k_p$, were already defecting at t_r according to the definition of \mathcal{A} . Hence,

$$n_{SD_i}^C(t_r+1) = n_{SD_i}^C(t_r) = n_{SD_i} \quad \forall i < k_p. \quad (14)$$

Now consider the SD_{k_p} agents. Let $\sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SD_{k_p}}^*$ be in force. Based on the definition of a^* , the following holds

$$n_{SD_{k_p}}^C(t_r) = 0. \quad (15)$$

On the other hand, according to (12),

$$n_C(t_r) \geq n_{SD_{k_p}}^* \Rightarrow \beta_{SD_{k_p}}(s_C(t_r)) = \{D\}.$$

Hence, due to (15),

$$n_{SD_{k_p}}^C(t_r+1) = n_{SD_{k_p}}^C(t_r) = 0.$$

In view of (13) and (14), the above equation results in

$$(\exists t_r : a(t_r) = a^*) \Rightarrow (a(t_r+1) = a^*).$$

Consequently, (10) can be proven by induction.

Now let $\sum_{j=1}^{k_p-1} n_{SD_j} < n_{SD_{k_p}}^*$ be in force. Then,

$$n_{SD_{k_p}}^C(t_r) \in \left\{ \lfloor n_{SD_{k_p}}^* \rfloor - \sum_{j=1}^{k_p-1} n_{SD_j}, \lceil n_{SD_{k_p}}^* \rceil - \sum_{j=1}^{k_p-1} n_{SD_j} \right\}.$$

If $n_{SD_{k_p}}^C(t_r) = \lfloor n_{SD_{k_p}}^* \rfloor - \sum_{j=1}^{k_p-1} n_{SD_j} = \lceil n_{SD_{k_p}}^* \rceil - \sum_{j=1}^{k_p-1} n_{SD_j}$, then $\lfloor n_{SD_{k_p}}^* \rfloor = \lceil n_{SD_{k_p}}^* \rceil = n_{SD_{k_p}}^*$ and hence $n_C(t_r) = n_{SD_{k_p}}^*$. Hence, the best reply for all of the SD_{k_p} agents at t_r is $\{C, D\}$. Hence, if an SD_{k_p} agent is chosen at t_r to update her strategy, she will not change her strategy. Hence,

$$\begin{aligned} n_{SD_{k_p}}^C(t_r+1) &= n_{SD_{k_p}}^C(t_r) \\ &= \lfloor n_{SD_{k_p}}^* \rfloor - \sum_{j=1}^{k_p-1} n_{SD_j} = \lceil n_{SD_{k_p}}^* \rceil - \sum_{j=1}^{k_p-1} n_{SD_j}. \end{aligned} \quad (16)$$

If $n_{SD_{k_p}}^C(t_r) = \lceil n_{SD_{k_p}}^* \rceil - \sum_{j=1}^{k_p-1} n_{SD_j}$ and $\lfloor n_{SD_{k_p}}^* \rfloor \neq n_{SD_{k_p}}^*$, then

$$n_C(t_r) = \lceil n_{SD_{k_p}}^* \rceil > n_{SD_{k_p}}^* \stackrel{(6)}{\Rightarrow} \beta_{SD_{k_p}}(s_C(t_r)) = \{D\}.$$

Hence, when an SD_{k_p} agent is chosen at t_r , we have that

$$n_{SD_{k_p}}^C(t_r+1) = n_{SD_{k_p}}^C(t_r) - 1 = \lfloor n_{SD_{k_p}}^* \rfloor - \sum_{j=1}^{k_p-1} n_{SD_j}. \quad (17)$$

If $n_{SD_{k_p}}^C(t_r) = \lfloor n_{SD_{k_p}}^* \rfloor - \sum_{j=1}^{k_p-1} n_{SD_j}$ and $\lceil n_{SD_{k_p}}^* \rceil \neq n_{SD_{k_p}}^*$, then

$$n_C(t_r) = \lfloor n_{SD_{k_p}}^* \rfloor < n_{SD_{k_p}}^* \stackrel{(6)}{\Rightarrow} \beta_{SD_{k_p}}(s_C(t_r)) = \{C\}.$$

Hence, when an SD_{k_p} agent is chosen at t_r , we have that

$$n_{SD_{k_p}}^C(t_r + 1) = n_{SD_{k_p}}^C(t_r) + 1 = \lceil n_{SD_{k_p}}^* \rceil - \sum_{j=1}^{k_p-1} n_{SD_j}. \quad (18)$$

Moreover, in general, when the chosen agent at t_r is not an SD_{k_p} agent, we have that

$$n_{SD_{k_p}}^C(t_r + 1) = n_{SD_{k_p}}^C(t_r). \quad (19)$$

Now (16), (17), (18) and (19) imply that

$$n_{SD_{k_p}}^C(t_r + 1) \in \left\{ \lfloor n_{SD_{k_p}}^* \rfloor - \sum_{j=1}^{k_p-1} n_{SD_j}, \lceil n_{SD_{k_p}}^* \rceil - \sum_{j=1}^{k_p-1} n_{SD_j} \right\}. \text{ Moreover, if } \sum_{j=1}^{k_p-1} n_{SD_j} < n_{SD_{k_p}}^*, \text{ then in case at some time } t,$$

In view of (13) and (14), the above equation results in

$$(\exists t_r : a(t_r) \in \{a_-^*, a_+^*\}) \Rightarrow (a(t_r + 1) \in \{a_-^*, a_+^*\}).$$

Consequently, (11) can be proven by induction, which completes the proof. \square

The time when the number of cooperators in the population equals that in \mathcal{A} , plays a key role in the rest of the proof. Hence, we proceed with the following definition. For $k_p \neq l + 1$, define \mathcal{B} as the set of the time steps t_b such that

$$\begin{cases} n_C(t_b) = \sum_{j=1}^{k_p-1} n_{SD_j} & \sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SD_{k_p}}^* \\ n_C(t_b) \in \{\lfloor n_{SD_{k_p}}^* \rfloor, \lceil n_{SD_{k_p}}^* \rceil\} & \sum_{j=1}^{k_p-1} n_{SD_j} < n_{SD_{k_p}}^* \end{cases}.$$

The following two lemmas are used to show that \mathcal{B} is infinite. In Lemma 3 it is shown that if the number of cooperators n_C exceeds the maximum number of cooperators in \mathcal{A} at some time t , it will return to the number of cooperators in \mathcal{A} after some finite time. The same is shown to happen in Lemma 4 when n_C becomes less than the minimum number of cooperators in \mathcal{A} .

Lemma 3: Let $k_p \neq l + 1$. If $\sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SD_{k_p}}^*$, then in case at some time t ,

$$n_C(t) > \sum_{j=1}^{k_p-1} n_{SD_j},$$

then a.s. there exists some natural number t_1 such that

$$n_C(t + t_1) = \sum_{j=1}^{k_p-1} n_{SD_j}.$$

Moreover, if $\sum_{j=1}^{k_p-1} n_{SD_j} < n_{SD_{k_p}}^*$, then in case at some time t ,

$$n_C(t) > \lceil n_{SD_{k_p}}^* \rceil,$$

then a.s. there exists some natural number t_1 such that

$$n_C(t + t_1) \in \{\lfloor n_{SD_{k_p}}^* \rfloor, \lceil n_{SD_{k_p}}^* \rceil\}.$$

Lemma 4: Let $k_p \neq l + 1$. If $\sum_{j=1}^{k_p-1} n_{SD_j} \geq n_{SD_{k_p}}^*$, then in case at some time t ,

$$n_C(t) < \sum_{j=1}^{k_p-1} n_{SD_j},$$

then a.s. there exists some natural number t_1 such that

$$n_C(t + t_1) = \sum_{j=1}^{k_p-1} n_{SD_j}.$$

Moreover, if $\sum_{j=1}^{k_p-1} n_{SD_j} < n_{SD_{k_p}}^*$, then in case at some time t ,

$$n_C(t) < \lfloor n_{SD_{k_p}}^* \rfloor,$$

then a.s. there exists some natural number t_1 such that

$$n_C(t + t_1) \in \{\lfloor n_{SD_{k_p}}^* \rfloor, \lceil n_{SD_{k_p}}^* \rceil\}.$$

According to Lemmas 3 and 4, the set \mathcal{B} is infinite. Now we define a function that is monotone in \mathcal{B} , and later use it to prove convergence to \mathcal{A} . First define $\alpha(t)$ as the greatest nonnegative number at t that satisfies the following conditions

$$n_{SD_j}^C(t) = n_{SD_j} \quad \forall j \in \{1, \dots, \alpha(t)\}, \quad (20)$$

$$\alpha(t) \leq k_p - 2. \quad (21)$$

Now define $h(t)$ as

$$h(t) := \sum_{i=1}^{\alpha(t)+1} n_{SD_i}^C(t). \quad (22)$$

Note that according to the definition, the maximum of h is $\sum_{j=1}^{k_p-1} n_{SD_j}$. The next two lemmas are required to show that $h(t)$ reaches its maximum value $\sum_{j=1}^{k_p-1} n_{SD_j}$.

Lemma 5: Let $k_p \neq l + 1$. For $t_b \in \mathcal{B}$, the following holds

$$h(t) \geq h(t_b) \quad \forall t \geq t_b. \quad (23)$$

Lemma 6: Let $k_p \neq l + 1$. If for some time $t_q \in \mathcal{B}$ it holds that

$$h(t_q) < \sum_{j=1}^{k_p-1} n_{SD_j}, \quad (24)$$

then a.s. there exists some time $t_r \in \mathcal{B}$, $t_r > t_q$ such that

$$h(t_r) > h(t_q). \quad (25)$$

Now the proof of Th. 1 can be done using the monotone property of h , which we avoid here due to space limitation.

After gaining insight into what the absorbing states are for the type population dynamics, as control engineers, we want to know whether we can shift the absorbing states and thus “control” the dynamics. In the next section, we present some of the ideas.

IV. CONTROLLING THE NUMBER OF COOPERATORS

Now that we know the behavior of the population game after a sufficiently long time, we are interested in investigating whether it is possible to control the number of cooperators in the final state. While there may be several ways to achieve this goal, e.g., adding some agents to the population, fixing the strategies of some of the agents to C or D , etc., perhaps the most natural parameter to control, which makes the least changes to the structure of the population game, is the type of each agent. Changing the type of an agent can be done by modifying the parameter(s) of her payoff matrix. By changing the types of a group of agents, a new type population is acquired which leads to a possibly new number of cooperators in the final state. Note that since the long run behavior of the population state is independent from the initial strategies of the agents, it does not matter when and in what order the changes in the types are done. For sake of simplicity we assume that they all happen at time $t = 0$. So the question we attack is how to set the number of cooperators in the long run to some reference value r by changing the types of a group of agents at some time step $t = 0$.

Each set of changes in the types of the agents results in some type population p . Moreover, if we know that type population p under which the number of cooperators a.s. equals the reference value for all time greater than some constant τ , we also know the type of which agent should be changed to what, by just comparing the original type population ξ with p . For $r \in \mathbb{Z}_{\geq 0}$ and a population of size n , define $\mathcal{F}(r, n)$ as the set of all feasible type-populations $p \in \mathcal{P}_n$ such that under the updater rule (2), the number of cooperators a.s. equals r for all time greater than some constant τ , i.e.,

$$\mathcal{F}(r, n) := \{p \mid p \in \mathcal{P}_n, (\exists \tau : n_C(p, t) = r \quad \forall t \geq \tau)\}.$$

Our goal is to determine $\mathcal{F}(r, n)$. Then it becomes clear how the types of the agents must change to have r cooperators in the long run. We need the following sets for a given $n \in \mathbb{N}$, $r \in \mathbb{Z}_{\geq 0}$ and $b \in \{1, \dots, l+1\}$:

$$\begin{aligned} \mathcal{F}_1^b(r, n) &:= \{p \mid p \in \mathcal{P}_n, k_p = b, \sum_{i=1}^{b-1} p_i = r\}, \\ \mathcal{F}_2^b(r, n) &:= \{p \mid p \in \mathcal{P}_n, k_p = b, \sum_{i=1}^{b-1} p_i < r\}. \end{aligned}$$

Theorem 2: Consider the types SD_1, \dots, SD_l and PD . Given $r \in \mathcal{D}_n$,

- 1) if there exists $b \in \{2, \dots, l\}$ such that $n_{SD_{b-1}}^* > r > n_{SD_b}^*$, then $\mathcal{F}(r, n) = \mathcal{F}_1^b(r, n)$;
- 2) if there exists $b \in \{1, \dots, l\}$ such that $r = n_{SD_b}^*$, then $\mathcal{F}(r, n) = \mathcal{F}_1^b(r, n) \cup \mathcal{F}_2^b(r, n)$;
- 3) if $r \in \{0, 1, \dots, \min\{\lceil n_{SD_l}^* \rceil - 1, n\}\}$, then $\mathcal{F}(r, n) = \mathcal{F}_1^{l+1}(r, n)$.

We skip the proof due to page limit.

V. CONCLUDING REMARKS

We have studied a finite heterogeneous population of game-playing agents under the myopic best response update rule. We have shown that based on the distribution of the types in the population, the total number of cooperators in the long run either reaches a certain steady state or fluctuates between two constants. Moreover, we have investigated how to control the number of cooperators, and in particular, for a given population and types, we show how to change the types of the agents to reach a desired number of cooperators in the long run.

REFERENCES

- [1] M. A. Nowak, "Five rules for the evolution of cooperation," *Science*, vol. 314, no. 5805, pp. 1560–1563, 2006.
- [2] P. A. Van Lange and J. A. Joireman, "How we can promote behavior that serves all of us in the future," *Social Issues and Policy Review*, vol. 2, no. 1, pp. 127–157, 2008.
- [3] F. Bullo, J. Cortes, and S. Martinez, *Distributed Control of Robotic Networks*. Princeton: Princeton University Press, 2009.
- [4] T. H. Chung, G. A. Hollinger, and V. Isler, "Search and pursuit-evasion in mobile robotics: A survey," *Autonomous Robots*, vol. 31, pp. 299–316, 2011.
- [5] N. Correll and A. Martinoli, "Multirobot inspection of industrial machinery," *IEEE Robotics and Automation Magazine*, vol. 16, pp. 103–112, 2009.
- [6] C. Hauert and M. Doebeli, "Spatial structure often inhibits the evolution of cooperation in the snowdrift game," *Nature*, vol. 428, no. 6983, pp. 643–646, 2004.
- [7] R. L. Riolo, M. D. Cohen, and R. Axelrod, "Evolution of cooperation without reciprocity," *Nature*, vol. 414, no. 6862, pp. 441–443, 2001.
- [8] O. Diekmann and S. van Gils, "On the cyclic replicator equation and the dynamics of semelparous populations," *SIAM Journal on Applied Dynamical Systems*, vol. 8, no. 3, pp. 1160–1189, 2009.
- [9] M. Zeeman and P. van den Driessche, "Three-dimensional competitive lotka-volterra systems with no periodic orbits," *SIAM Journal on Applied Mathematics*, vol. 58, no. 1, pp. 227–234, 1998.
- [10] P. Ramazi and M. Cao, "Stability analysis for replicator dynamics of evolutionary snowdrift games," *In Proc. of the 53rd IEEE Conference on Decision and Control*, pp. 4515–4520, 2014.
- [11] W.-B. Du, X.-B. Cao, M.-B. Hu, and W.-X. Wang, "Asymmetric cost in snowdrift game on scale-free networks," *EPL (Europhysics Letters)*, vol. 87, no. 6, p. 60004, 2009.
- [12] J. Dijkstra and M. A. van Assen, "Network public goods with asymmetric information about cooperation preferences and network degree," *Social Networks*, vol. 35, no. 4, pp. 573–582, 2013.
- [13] R. Waldeck, "Segregated cooperation," *Journal of Artificial Societies and Social Simulation*, vol. 16, no. 4, p. 14, 2013.
- [14] G. E. Kreindler and H. P. Young, "Fast convergence in evolutionary equilibrium selection," *Games and Economic Behavior*, vol. 80, pp. 39–67, 2013.
- [15] C. Alós-Ferrer, "Finite population dynamics and mixed equilibria," *International Game Theory Review*, vol. 5, no. 03, pp. 263–290, 2003.
- [16] C. Alós-Ferrer and N. Netzer, "The logit-response dynamics," *Games and Economic Behavior*, vol. 68, no. 2, pp. 413–427, 2010.
- [17] J. Riehl and M. Cao, "Towards control of evolutionary games on networks," *In Proc. of the 53rd IEEE Conference on Decision and Control*, pp. 2877–2882, 2014.
- [18] —, "Towards optimal control of evolutionary games on networks," *IEEE Transactions on Automatic Control*, 2015, under review.
- [19] H. Liang, M. Cao, and X. Wang, "Analysis and shifting of stochastically stable equilibria for evolutionary snowdrift games," *Systems & Control Letters*, vol. 85, no. 3, pp. 16–22, 2015.
- [20] A. Yanase, "Dynamic games of environmental policy in a global economy: taxes versus quotas," *Review of International Economics*, vol. 15, no. 3, pp. 592–611, 2007.
- [21] P. Ramazi, J. Hessel, and M. Cao, "How feeling betrayed affects cooperation," *PLoS ONE*, vol. 10, no. 4, 2015.